#4 a
s/2 19 020477

Europäisches Patentamt | European Patent Office | Office européen des brevets

J1046 U.S. PTO
10/046632
01/14/02

Bescheinigung Certificate Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

01200143.4

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

I.L.C. HATTEN-HECKMAN

DEN HAAG, DEN
THE HAGUE, 25/10/01
LA HAYE, LE

EPA/EPO/OEB Form 1014 - 02.91

Eur päisches Patentamt | European Patent Office | Office eur péen des brevets

Blatt 2 d r Bescheinigung
Sheet 2 of the certificate
Page 2 de l'attestation

Anmeldung Nr.:
Application no.: 01200143.4
Demande n°:

Anmeldetag:
Date of filing: 16/01/01
Date de dépôt:

Anmelder:
Applicant(s):
Demandeur(s):
Koninklijke Philips Electronics N.V.
5621 BA Eindhoven
NETHERLANDS

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:
Sinusoidal coding

In Anspruch genommene Prioriät(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat: State: Pays:	Tag: Date: Date:	Aktenzeichen: File no. Numéro de dépôt:

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:
/

Am Anmeldetag benannte Vertragstaaten:
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE/TR
Etats contractants désignés lors du depôt:

Bemerkungen:
Remarks:
Remarques:

Frequency warping in sinusoidal coding

Bert den Brinker

Audio coding schemes which are based on sinusoidal coding employ a certain segment size, or multiple segment sizes (multiscale models) for the estimation of the sinusoidal parameters and the extraction of the associated components. In a one-scale model, the time-frequency resolution trade-off setting is a major determinant in the final quality bit-rate compromise while, for multiscale models, problems arise due to the scattering of components over scales and the consequent parameter merging process. To overcome these problems it is proposed to use a single-scale frequency-warped sinusoidal estimation mechanism where the warped frequency scale resembles that of the human ear.

Things known so far

The common approach in sinusoidal coding of audio is to segment the signal and estimate the sinusoids within that segment, e.g. [1, 2, 3, 4, 5]. This gives problems with the required time-frequency resolution trade-off, especially for high-quality audio coding where a large frequency range is necessary. Therefore, multiscale models have been proposed [6, 7], but these bring about the problem of scattering of components over scales and/or of merging the data retrieved at different scales.

For LPC-coding of audio, it has been suggested to work in the warped frequency domain [8, 9, 10, 11, 12], where the warping is related to perceptually relevant scales, e.g., the Bark or ERB scale [13, 14].

The problem for which this invention brings the solution

The problem is either the time-frequency resolution in single scale model or the co-operation of the different scales (realised in parallel paths) in a multi-scale sinusoidal estimation/extraction mechanism.

Proposed measures

In order to obtain a sinusoidal estimation mechanism with a time-frequency resolution close to that employed in the human ear, it is proposed to use the warping technique.

Embodiment

In current sinusoidal coders, a tapped-delay-line is used to define a segment of data which is input to the sinusoidal analysis module. Next, this data is analysed for sinusoidal content, typically the data is windowed and Fourier transformed to detect the relevant sinusoidal components.

Instead of a tapped-delay-line, a first-order allpass section can be used to replace each delay. Taking the Fourier transform of the outputs of the allpass section, we obtain the Fourier transform on a frequency-warped scale. The sinusoidal extraction can be done as usual.

After sinusoidal extraction, the subsequent processing stage is residual modelling. The cheapest way of residual modelling is probably using a parametric model for the power spectral density functions. We note here that such an approach allows the integration of sinusoidal and noise estimation since, for noise modelling, warped LPC [8, 9, 10, 11, 12] or warped ARMA modelling (according to PHNL000287 and PHNL000288) can be used.

Application areas

Audio and speech coding.

References

[1] R. McAulay and T. Quartieri. Speech analysis/synthesis based on sinusoidal representation. *IEEE Trans. Acoust., Speech, Signal Process.*, 43:744–754, 1986.

[2] B. Edler, H. Purnhagen, and C. Ferekidis. ASAC – Analysis/synthesis codec for very low bit rates. Preprint 4179 (F-6) 100th AES Convention, Copenhagen, 11-14 May 1996.

[3] K.N. Hamdy, M. Ali, and A.H. Tewfik. Low bit rate high quality audio coding with combined harmonic and wavelet representations. In *Proc. 1996 Int. Conf. Acoust. Speech Signal Process. (ICASSP96)*, pages 1045–1048, Atlanta GA, 7-10 May 1996. IEEE, Picataway, NJ.

[4] P. Masri. *Computer modelling of sound for transformation and synthesis of musical signals.* PhD thesis, Univ. of Bristol, 1996.

[5] X. Serra. Musical sound modeling with sinusoids plus noise. In C. Roads, S. Pope, A. Picialli, and G. De Poli, editors, *Musical Signal Processing*. Swets & Zeitlinger, 1997.

[6] M.M. Goodwin. *Adaptive signal models: theory, algorithms, and audio applications.* PhD thesis, Univ. of California, Berkeley, 1997. Pages 1–259.

[7] S.N. Levine. *Audio representation for data compression and compressed domain processing.* PhD thesis, Stanford Univ. (CA), 1999. Pages 1–136.

[8] H.W. Strube. Linear prediction on a warped frequency scale. *J. Acoust. Soc. Am.*, 68:1071–1076, 1980.

[9] U.K. Laine, M. Karjalainen, and T. Altosaar. WLP in speech and audio processing. In *Proc. ICASSP-94*, volume II, pages 349–352, Adelaide, Australia, 1994.

[10] A. Härmä, U.K. Laine, and M. Karjalainen. WLPAC – A perceptual audio codec in a nutshell. In *Proc. AES*, pages XX–YY, 96.

[11] A. Härmä, U.K. Laine, and M. Karjalainen. Warped linear prediction (WLP) in audio coding. In *Proc. NorSig-96*, pages 447–450, Espoo, Finland, 24-27 Sept. 1996.

[12] A. Härmä, U.K. Laine, and M. Karjalainen. An experimental audio codec based on warped linear prediction of complex valued signals. In *Proc. ICASSP-97*, pages 323–326, Munich, Germany, 21-24 April 1997.

[13] E. Zwicker and H. Fastl. *Psychoacoustics: Facts and Models.* Springer, Berlin, 1990.

[14] B.C.J. Moore. *An introduction to the psychology of hearing.* Academic Press, San Diego (CA), 1997.

PHNL 000287 = non pre-published PCT patent application nr. EP 00/04599 filed 17.05.00

PHNL 000288 = non pre-published PCT patent application nr. EP00/04601 filed 17.05.00

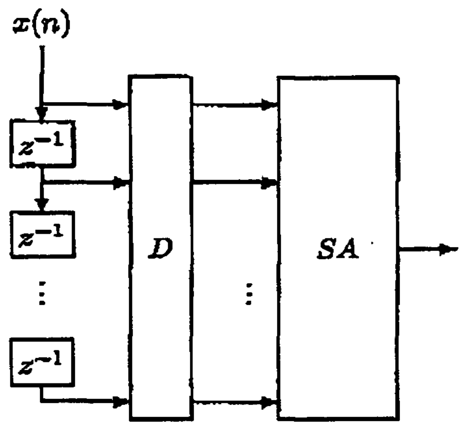

Figure 1: Usual sinusoidal analysis. The signal x is input to a delay line. The content of the delay line is input to a downsampler: once per D samples the input is passed to a the sinusoidal analysis SA. Typically, the consecutive segments that are input to SA have overlap. The output of the mechanism SA are the sinusoidal parameters describing the signal.

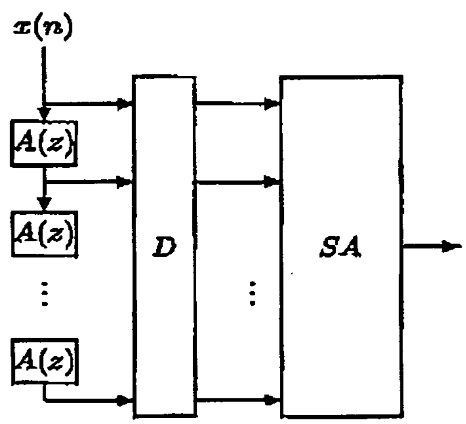

Figure 2: Warped Fourier transform by using allpass section $A(z)$ followed by a downsampler D. The output samples are input to mechanism for sinusoidal analysis SA. The output of this mechanism are the sinusoidal parameters describing the signal. In view of the warping operation, a pre-filtering operation can be applied to x for (partial) amplitude and/or amplitude phase-compensation of the allpass line.

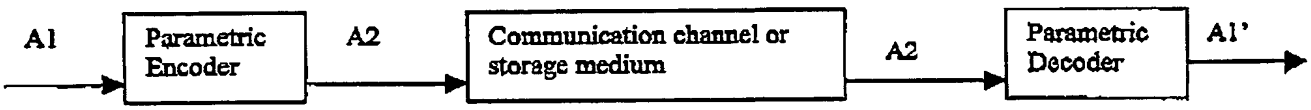

Fig. 3

Fig. 3 shows an embodiment of the invention. An audio and/ or speech signal A1 is furnished to a parametric encoder and coded into an encoded audio and/or speech signal A2. The encoded signal A2 is transmitted over a communication channel or stored on a storage medium. A parametric decoder obtains the encoded signal from the communication channel or storage medium and decodes this signal A2 into a decoded audio and/or speech signal A1' which is a representation of A1. The parametric encoder according to this embodiment of the invention estimates sinusoidal parameters on a frequency warped scale. The estimated sinusoidal parameters are included in the bit-stream A2 and transmitted to the decoder. In the decoder, on the basis of these sinusoidal parameters which have been estimated on a frequency warped scale, a reconstruction of the original audio signal is made: A1'.

CLAIMS:

1. A parametric coding method of encoding an audio (and/ or speech) signal, which method comprises the steps of estimating sinusoidal parameters and extracting associated components, wherein the estimating step is performed on a frequency warped scale.

2. A parametric encoder for encoding an audio (and/ or speech) signal, which device comprises means for estimating sinusoidal parameters and extracting associated components, wherein the estimating step is performed on a frequency warped scale.

3. A parametric decoding method of decoding an encoded audio (and/ or speech) signal, which method comprises the step of receiving the encoded audio signal which includes sinusoidal parameters which have been estimated on a frequency warped scale, and using said sinusoidal parameters in the reconstruction of an audio signal.

4. A parametric decoder for decoding an encoded audio and/ or speech signal, which decoder comprises means for receiving the encoded audio signal which includes sinusoidal parameters which have been estimated on a frequency warped scale, and means for using said sinusoidal parameters in the reconstruction of an audio signal.

5. An encoded audio and/ or speech signal, which signal includes sinusoidal parameters which have been estimated on a frequency warped scale